H. B. Quinn Code y

Technical Report 70-117
GJ-231 and NGL-21-002-008

June 1970

On Classes of n-dimensional Nonlinear Mappings Generalizing Several Types of Matrices

Werner C. Rheinboldt
Research Professor

UNIVERSITY OF MARYLAND
COMPUTER SCIENCE CENTER
COLLEGE PARK, MARYLAND

FACILITY FORM 602

N 71-70563
(ACCESSION NUMBER)

(THRU)

43
(PAGES)

(CODE)

CR-116166
(NASA CR OR TMX OR AD NUMBER)

19
(CATEGORY)

RECEIVED
NASA STI FACILITY
INPUT BRANCH
FEB 1971

Technical Report 70-117
GJ-231 and NGL-21-002-008

June 1970

On Classes of n-dimensional Nonlinear Mappings Generalizing Several Types of Matrices

Werner C. Rheinboldt
Research Professor

Invited Address at the Symposium on the Numerical Solution of Partial Differential Equations II, University of Maryland, May 11-15, 1970. To be published in the Proceedings of that Conference by Academic Press, Inc., New York, N.Y.

This work was supported in part by the National Science Foundation under Grant GJ-231 and the National Aeronautics and Space Administration under Grant NGL-21-002-008.

Abstract

This paper concerns several related classes of mappings from R^n into itself which represent nonlinear generalizations of certain types of matrices, including the diagonally dominant and the monotonic matrices, as well as the M-, P-, and S-matrices and their weaker forms. Analogous to the linear case, these nonlinear mappings occur frequently as discrete analogs of boundary value problems and in network flow problems.

The different function classes have been introduced and analyzed in recent work by J. Moré and W. Rheinboldt. This article begins with a survey of that work which covers, in particular, the basic definitions of these mappings, their principal properties, as well as their interrelations. Then several results are proved concerning surjectivity properties of some of the functions, thereby generalizing in part various older results. The relation between nonlinear mappings and equilibrium problems for network flows is discussed, and it is shown how some of the properties of the functions under consideration lead to statements about Dirichlet-type problems for network flows and about a general maximum principle. Finally, a class of implicit iterative processes is introduced which represents a generalization of the family of linear methods obtained from regular splittings. For these processes both a monotonic and a global convergence theorem are proved. As an application, this ensures the global convergence of the (underrelaxed) block-Jacobi and block-Gauss-Seidel methods for continuous, surjective M-functions, which, in turn, generalizes a corresponding theorem of Rheinboldt for the point processes.

On Classes of n-dimensional Nonlinear Mappings Generalizing Several Types of Matrices[1)]

by

Werner C. Rheinboldt[2)]

1. Introduction

Consider a nonlinear mapping $F:D \subset R^n \to R^n$ and the system of equations $Fx = z$. Many of the well-known convergence results about iterative processes for solving this system place only very general analytic conditions upon F, such as differentiability, Lipschitz-continuity, etc. This provides, of course, for rather broad theorems which are often generalizable to infinite-dimensional spaces. But at the same time, when applied to particular mappings on R^n, such as, for example, discrete analogs of elliptic boundary value problems, or nonlinear network flow functions, these general convergence results tend to give only relatively limited or localized information.

The situation is analogous to the one in which only a norm condition $\|B\| < 1$ is used to ensure the convergence of an iterative process $x^{k+1} = Bx^k + z$, $k = 0,1,\ldots$, for solving the linear system $Ax = z$. It is well-known that stronger convergence results here require a much deeper knowledge of the spectral properties of the iteration matrix B and hence of the structural properties of A itself. Similarly, it appears to be beyond question that also in the nonlinear case stronger convergence

---

1) This work was supported in part by the National Science Foundation under Grant GJ-231 and the National Aeronautics and Space Administration under Grant NGL-21-002-008.

2) Computer Science Center, University of Maryland, College Park, Md. 20742.

theorems will have to be based, in general, on more specific assumptions about the inherent finite-dimensional structure of the mapping F, as, for instance, the specific dependence of the components f_j of F on the individual variables x_i. This in turn leads to the need for defining and analyzing appropriate classes of n-dimensional nonlinear mappings as they occur in various applications. So far only a few structurally different classes of such mappings have been considered, and a need for more work along this line certainly exists.

This article presents a survey--and some new results--on recent work about a group of related classes of n-dimensional functions. Following an unpublished suggestion of Ortega, Rheinboldt [1969b] investigated the so-called M-functions on R^n which represent a nonlinear generalization of the well-known M-matrices. In particular, it was shown that the discrete analogs of mildly nonlinear elliptic problems considered by Bers [1953], Greenspan and Parter [1965], Ortega and Rheinboldt [1967], [1970a] and others, as well as the network flow functions analyzed by Birkhoff and Kellogg [1966] and Porsching [1969], are specific cases of M-functions. Moreover, the global convergence of the (underrelaxed) nonlinear (point-) Jacobi-, and (point-) Gauss-Seidel processes was established for continuous, surjective M-functions, thereby generalizing the corresponding well-known results for M-matrices (see, e.g., Varga [1962]).

The latter result is typical for many surprising similarities between the behavior of M-functions and M-matrices, and in turn these similarities suggest the idea of looking for analogous nonlinear extensions of other types of matrices as well. In this connection, the P- and S-matrices and their weaker forms, the P_0- and S_0-matrices considered by Fiedler and Ptak [1962], [1966] are of some interest, especially since every M-matrix is also a P-, as well as an S-matrix. Recently, Moré and Rheinboldt [1970] introduced and studied such nonlinear n-dimensional generalizations of these four types of matrices--accordingly named P_0, P, S_0, and S-functions. As expected, M-functions are special cases of P-functions and the continuous P-functions are S-functions. It also turned out that earlier results of Gale and Nikaido [1965] and Karamardian [1968] have a natural place in this theory, and that certain mappings, considered by Willson [1968] and Sandberg and Willson [1969a/b] in connection with particular electronic circuit problems, are included among these new functions.

In Section 2 we present the basic definitions of the mentioned function-classes and of several related types of mappings. This is followed in Section 3 by a survey of the major properties of these functions and of their interrelations. For clarity the results are not always stated in their most general form, and for further details about the material in the first two sections, as well as for many of the proofs, reference is made to Rheinboldt [1969b], Moré and Rheinboldt [1970], and Moré [1970]. Section 4 concerns the problem of determining the surjectivity of certain of the mappings under consideration and presents some new generalizations of earlier results on M-functions. In Section 5 connections between n-dimensional nonlinear mappings and network flow problems are discussed, and, finally Section 6 concerns a general type of iterative process similar to the processes obtained by regular splittings in the linear case. In particular, a global convergence theorem is proved which covers as a special case the convergence of the block-Jacobi-, and block-Gauss-Seidel process for continuous surjective M-functions.

At this point, I would like to extend my special thanks to Jorge Moré for his helpful cooperation in preparing this article and to the Gesellschaft für Mathematik und Datenverarbeitung, m.b.H., Birlinghoven/Germany, where, in 1969, I began work on several of the new results reported here.

2. Basic Definitions

Throughout this paper, $x \leq y$ denotes the natural (component-wise) partial ordering on the n-dimensional real linear space R^n of column vectors, and $x < y$ stands for $x_i < y_i$, $i \in N = \{1,2,\ldots,n\}$. The corresponding notation is used on the space $L(R^n)$ of real $n \times n$ matrices.

We begin by recalling the following standard terminology:

Definition 2.1. (a) A mapping $F:D \subset R^n \to R^n$ is isotone (or antitone) on D if $x \leq y$, $x,y \in D$, implies that $Fx \leq Fy$ (or $Fx \geq Fy$), and strictly isotone (or strictly antitone) if, in addition, it follows from $x < y$, $x,y \in D$, that also $Fx < Fy$ (or $Fx > Fy$).

(b) The function $F:D \subset R^n \to R^n$ is inverse isotone on D if $Fx \leq Fy$, $x,y \in D$, implies that $x \leq y$.

Note the self-evident fact that an affine mapping $Fx = Ax + b$ is isotone exactly if $A \geq 0$ and inverse isotone if and only if A is nonsingular and $A^{-1} \geq 0$.

There is a close connection between nonlinear network flows and several of the function classes to be discussed here. In fact, many of the results about these functions have inherent network-theoretical aspects and appear to be intuitively clearer if a network terminology is used to state them. Following Rheinboldt [1969b]--and in analogy with the connection between graphs and their incidence matrices--a particular network is associated with any function on R^n.

<u>Definition 2.2</u>. Consider $F:D \subset R^n \to R^n$ with the components $f_1,\ldots,f_n$.

(a) For any fixed $x \in R^n$ the n^2 functions

$$\psi_{ij}: \{t \in R^1 \mid x+te^j \in D\} \to R^1,\ \psi_{ij}(t) = f_i(x+te^j),\ i,j \in N$$

are the <u>link-functions</u> of F at x. Here e^j are the usual unit basis vectors in R^n.

(b) The <u>associated network</u> $\Omega_F = \{N,\Lambda_F\}$ of F consists of the set of nodes $N = \{1,\ldots,n\}$ and the set of links

$$\Lambda_F = \{(i,j) \in N \times N \mid i \neq j,\ \psi_{ij} \text{ not constant for some } x \in R^n\}.$$

A link $(i,j) \in \Lambda_F$ is <u>permanent</u> if ψ_{ij} is not constant for any $x \in R^n$.

This notation can be interpreted as follows: The variables $x_1,\ldots,x_n$ are state variables associated with the n nodes of Ω_F, and the value $f_i(x_1,\ldots,x_n)$ is the (total) efflux from node i at state x. The nodes i and j of N are connected by a link, if there is at least one state x at which the link function ψ_{ij} is not constant; we might say that at state x the link $(i,j) \in \Lambda_F$ is conducting. A permanent link is then conducting at any state x.

We shall now place various conditions upon the behavior of the link functions, and in all cases these conditions will be assumed permanent, that is, they are to hold independently of the particular state.

Definition 2.3. A mapping $F:D \subset R^n \to R^n$ is off-diagonally antitone if for any state $x \in R^n$ the "off-diagonal" link functions ψ_{ij}, $i \neq j$, $i,j \in N$, are antitone. Similarly, F is diagonally (strictly) isotone if for any $x \in R^n$ the "diagonal" link functions $\psi_{11},\ldots,\psi_{nn}$ are (strictly) isotone.

Off-diagonal antitonicity states that for any linked nodes $i,j \in N$ a change of the state x_j of the receiving node produces a change with the opposite sign in the efflux f_i from the originating node i. This is, of course, the expected situation in a linear potential network where the flow from i to j is proportional to the potential difference $x_i - x_j$. The matrix $A = (a_{ij})$ describing such a linear network flow then satisfies $a_{ij} \leq 0$, $i \neq j$, which is one of the properties of an M-matrix. Since the other property, $A^{-1} \geq 0$, is equivalent with inverse isotonicity, we are led to the following nonlinear generalization of M-matrices.

Definition 2.4. An inverse isotone and off-diagonally antitone mapping $F:D \subset R^n \to R^n$ is an M-function.

It is now hardly surprising that an affine mapping $Fx = Ax + b$ is an M-function if and only if $A \in L(R^n)$ is an M-matrix.

Ky Fan [1958] has shown that all principal minor determinants of an M-matrix are necessarily positive. The same result, of course, holds for all symmetric, positive definite matrices. More generally, Fiedler and Ptak [1962] considered the class of all matrices in $L(R^n)$ with this property and called them P-matrices.

A different generalization of the M-matrices can be obtained from the following characterization result of Ky Fan [1958]: A matrix $A \in L(R^n)$ with $a_{ij} \leq 0$, $i \neq j$, $i,j \in N$, is an M-matrix if and only if $Au > 0$ for some $u > 0$. Following earlier work by Stiemke [1915], Fiedler and Ptak [1966] called any $A \in L(R^n)$ an S-matrix if $Au > 0$ for some $u > 0$. In the same article they also showed that any P-matrix is an S-matrix and proved a number of results about these and related matrices.

Stimulated by these linear results, as well as by some nonlinear results of Gale and Nikaido [1965], Karamardian [1968], and Sandberg and Willson [1969a], Moré and Rheinboldt [1970] introduced the following nonlinear generalizations of the P- and S-matrices and of their weaker forms:

<u>Definition 2.5</u>. (a) A mapping $F:D \subset R^n \to R^n$ is a <u>P_0-function</u> (or <u>P-function</u>) on D, if for any $x,y \in D$, $x \neq y$, there exists a $k \in N$ such that $(x_k-y_k)(f_k(x)-f_k(y)) \geq 0$, $x_k \neq y_k$, (or $(x_k-y_k)(f_k(x)-f_k(y)) > 0$).

(b) $F:D \subset R^n \to R^n$ is an <u>S_0-function</u> (or <u>S-function</u>) on D, if for any $x \in D$ there exists a $y \in D$ such that $y \geq x$, $y \neq x$, and $Fy \geq Fx$ (or $Fy > Fx$).

Again it is easily verified--using the results of Fiedler and Ptak [1962], [1966]--that an affine mapping $Fx = Ax + b$ belongs to one of these four classes of functions if and only if $A \in L(R^n)$ is a member of the corresponding class of matrices.

In network terminology, P-functions have the property that for any (non-zero) change of the state there is at least one node at which the change of the efflux has the same sign as the change of state. For many applications this appears to be a rather natural condition.

If $A \in L(R^n)$ is diagonally dominant and has a non-negative diagonal, then A is a P_0-matrix, since, if $x \neq 0$ and $k \in N$ such that $|x_k| = \|x\|_\infty$, we have

$$x_k(Ax)_k \geq x_k^2(a_{kk} - \sum_{j \neq k} |a_{kj}|) \geq 0, \; x_k \neq 0.$$

If A is even strictly diagonally dominant, it is a P-matrix. This suggests the question whether the concept of diagonal dominance can also be extended to nonlinear functions. Already a simple reflection shows that several natural direct generalizations are not entirely satisfactory; this makes the following ingenious definition of Moré [1970] rather interesting:

Definition 2.6. A mapping $F:D \subset R^n \to R^n$ is strictly diagonally dominant, if for any $x,y \in D$, $x \neq y$, it follows from $f_k(x) = f_k(y)$ that $|x_k-y_k| < \|x-y\|_\infty$.

Moré [1970] shows that, again, an affine mapping $Fx = Ax + b$ is strictly diagonally dominant if and only if A is a strictly diagonally dominant matrix. He also introduces an extension of the concept which includes the irreducibly diagonally dominant matrices. This generalization is based on the existence of certain paths in the associated network.

3. Properties of the Different Functions

In line with the survey nature of this article we summarize now without proof some of the major properties of the classes of functions introduced in the previous section. For the sake of simplicity, these results are not given in their most general form, and, in particular, it is always assumed that $F:R^n \to R^n$ is defined on all of R^n, although more restricted domains could also be admitted.

Theorem 3.1 - Relations between the Classes.

(a) Any P- or S-function $F:R^n \to R^n$ is also a P_0- or S_0-function, respectively.

(b) Any continuous P_0 or P-function $F:R^n \to R^n$ is also an S_0 or S-function, respectively.

(c) Any isotone, or antitone mapping $F:R^n \to R^n$ is an S_0-function, and strictness implies that F is an S-function.

(d) A continuous, inverse-isotone mapping $F:R^n \to R^n$ is an S-function.

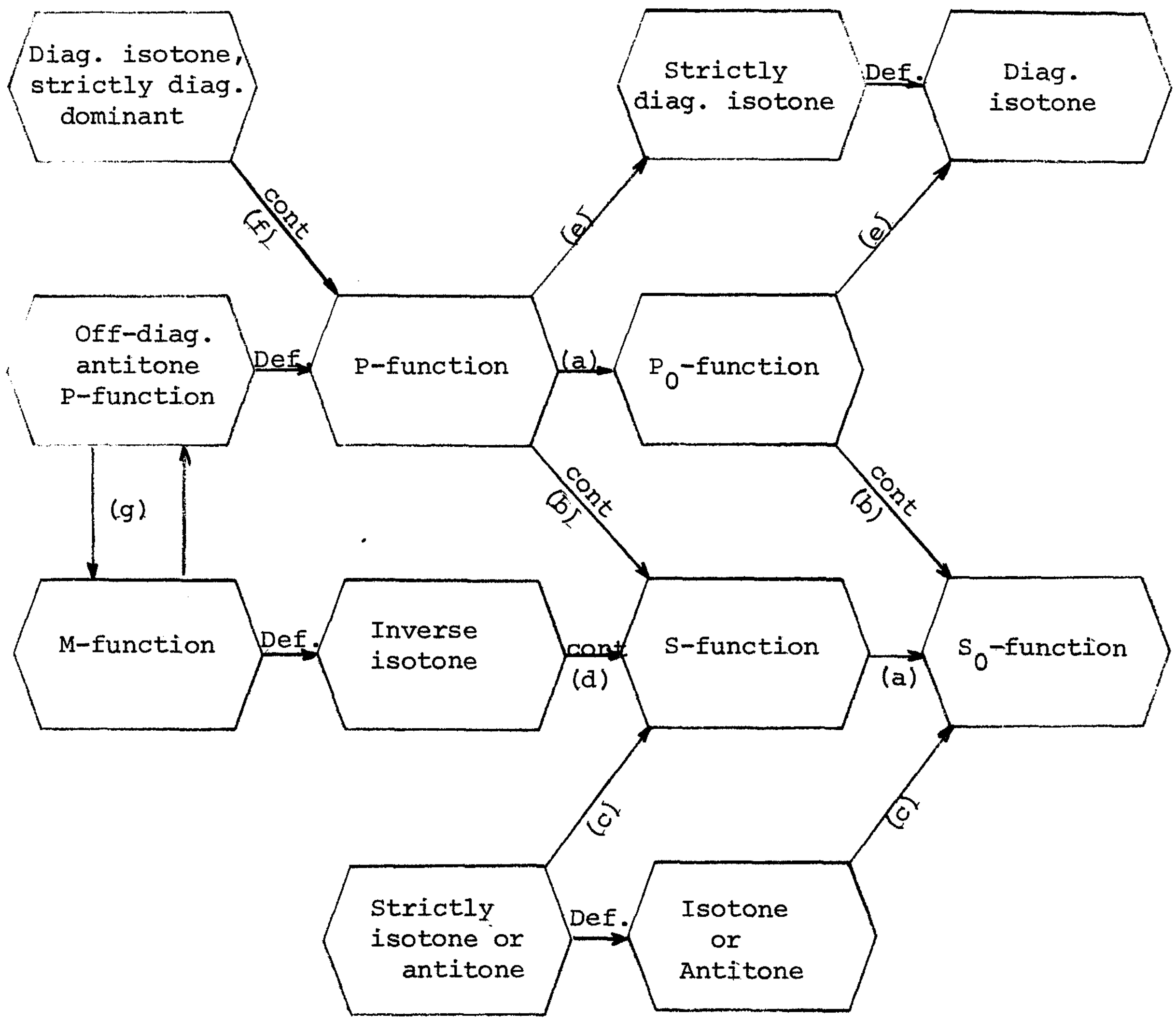
(e) If $F:R^n \to R^n$ is a P_0- or P-function, then F is diagonally isotone, or strictly diagonally isotone, respectively.

(f) Any continuous, diagonally isotone, and strictly diagonally dominant mapping $F:R^n \to R^n$ is a P-function.

(g) $F:R^n \to R^n$ is an M-function if and only if it is an off-diagonally antitone P-function.

The implications (a), (c), and (e) are rather straightforward consequences of the definitions; (b) represents a result of Karamardian [1968] phrased in this terminology; (d) and (g) are proved by Moré and Rheinboldt [1970], while (f) is a result of Moré [1970].

We indicate the general structure of these relations in the following diagram:



Note that, by definition, any isotone, or strictly isotone function is diagonally isotone, or strictly diagonally isotone, respectively. This relation is not shown. Note also that the diagram contains several derived implications, such as, for example, that any M-function is strictly diagonally isotone.

Theorem 3.2 - Inverses.

(a) $F:R^n \to R^n$ is inverse isotone if and only if F is injective and $F^{-1}:FR^n \to R^n$ is isotone.

(b) If $F:R^n \to R^n$ is a P-function, then F is injective and $F^{-1}:FR^n \to R^n$ is again a P-function.

(c) If $F:R^n \to R^n$ is an F-differentiable, injective P_0-function, then $F^{-1}:FR^n \to R^n$ is again a P_0-function.

The proofs of (a) and (b) are straightforward consequences of the definitions. Part (c) is proved by Moré and Rheinboldt [1970]; it is conjectured that the result remains valid if F is only continuous.

In the case of P_0-, P-, or M-matrices also any principal submatrix belongs to the same class. In order to consider the nonlinear analog of this result, we formalize first the concept of a subfunction.

Definition 3.3. The subfunction of the mapping $F:R^n \to R^n$ corresponding to the index set $M = \{i_1,\dots,i_m\} \subset N$, $0 < m \leq n$, and the constants c_j, $j \notin M$, is the mapping $G:R^m \to R^m$ with the components

$$g_k(y) = f_{i_k}\Big(\sum_{j=1}^{m} y_j e^{i_j} + \sum_{j \notin M} c_j e^j\Big),\ k=1,\dots,m,\ y \in R^m,$$

where e^j are the unit basis vectors in R^n.

We shall see in Section 5 that these subfunctions are of particular interest in connection with Dirichlet boundary value problems for network flows.

In generalization of the cited result for P_0-, P-, and M-matrices we have now:

Theorem 3.4 - Subfunctions. For any P_0-, P-, or M-function $F:R^n \to R^n$, also any subfunction of F belongs to the same function class.

In the case of P_0- and P-functions, the proof follows directly from the definitions. For continuous, surjective M-functions, the result was given by Rheinboldt [1969b]. Its generalization to arbitrary M-functions is due to Moré and Rheinboldt [1970]; interestingly, the proof is based on the corresponding result for P-functions together with the relation (g) of Theorem 3.1 between the two function classes.

In the applications, characterization theorems for the functions of the various classes are of considerable importance. For M-functions Rheinboldt [1969b] gave four related theorems of this type, none of which required more than continuity. In order to illustrate the connection to other results given later, we prove here some modification of one of these theorems. As usual, a path from i to ℓ in Ω_F is a sequence of links of the form

$$(3.1) \qquad (i_0,i_1),\ (i_1,i_2),\ldots,(i_m,i_{m+1}),\ i_0 = i,\ i_{m+1} = \ell,\ m \geq 0.$$

Theorem 3.5. Let $F:R^n \to R^n$ be off-diagonally antitone, and assume that for any $x \in R^n$ there exists a vector $u = u(x) > 0$ such that the mapping

$$p^x : R^1 \to R^n,\ p_i^x(t) = f_i(x+tu(x)),\ i \in N,$$

is isotone. Suppose further that for any $x \in R^n$ and $i \in N$ there is a path (3.1) to a node $\ell = \ell(i,x) \in N$ such that p_ℓ^x is strictly isotone and that at any state the link functions $\psi_{i_j i_{j+1}}$, $j = 0,\ldots,m$, are strictly antitone. Then F is an M-function.

<u>Proof</u>: Suppose that $Fx \leqslant Fy$. Then, with $u = u(y)$,

$$+\infty > t_0 = \inf\{t \in R^1 \mid tu \geqslant y-x\} > -\infty$$

and $N_0 = \{i \in N \mid t_0 u_i = y_i - x_i\}$ is not empty. If $t_0 \leqslant 0$, then $x \leqslant y$, hence suppose that $t_0 > 0$. If $i \in N_0$, then also $j \in N_0$ for any $j \neq i$ such that ψ_{ij} is always strictly antitone. In fact, otherwise $y_i + t_0 u_i = x_i$ and $y_j + t_0 u_j > x_j$, and

$$f_i(y+t_0u) < f_i(y_1+t_0u_1,\ldots,y_{j-1}+t_0u_{j-1},x_j,y_{j+1}+t_0u_{j+1},\ldots,y_n+t_0u_n)$$

$$\leqslant f_i(x) \leqslant f_i(y) \leqslant f_i(y+t_0u)$$

provides a contradiction. Hence, there exists a node $i \in N_0$ such that p_i^y is strictly isotone. But then

$$f_i(y+t_0u) = f_i(y_1+t_0u_1,\ldots,y_{i-1}+t_0u_{i-1},x_i,y_{i+1}+t_0u_{i+1},\ldots,y_n+t_0u_n)$$

$$\leqslant f_i(x) \leqslant f_i(y) < f_i(y+t_0u)$$

is again a contradiction. Altogether, therefore, $t_0 > 0$ is impossible and the result is proved.

Note that when P^x is strictly isotone for any x, then F is already an M-function since we can take $\ell = \ell(i,x) = i$. This generalizes the sufficiency portion of the earlier cited characterization of M-matrices by Ky Fan, namely, that A is an M-matrix if and only if $a_{ij} \leq 0$, $i \neq j$, and $Au > 0$ for some $u > 0$. The necessity part is trivial, since we can take $u = A^{-1}e > 0$, with $e = (1,1,\ldots,1)^T$. It might be conjectured that similarly for M-functions the conditions of Theorem 3.5 are also necessary. This is not the case as the following example shows:

$$F: R^2 \to R^2, \quad Fx = \begin{pmatrix} (\arctan x_1) - x_2 \\ \arctan x_2 \end{pmatrix} .$$

It is readily verified that F is an M-function, but for no $u > 0$ is $F(tu)$ an isotone function of t for all $t \in R^1$.

The above characterization result for M-functions does not even require F to be continuous. If F is assumed to be differentiable, simpler characterizations can be obtained in terms of properties of the derivative.

<u>Theorem 3.6</u>. Let $F: R^n \to R^n$ be F-differentiable on all of R^n.

(a) F is a P_0-function if and only if, for any $x \in R^n$, $F'(x)$ is a P_0-matrix.

(b) If, for any $x \in R^n$, $F'(x)$ is a P-matrix, then F is a P-function.

(c) If, for any $x \in R^n$, $F'(x)$ is an M-matrix, then F is an M-function.

(d) If F is an M-function, then $F'(x)$ is an M-matrix whenever it is nonsingular.

(e) If, for any $x \in R^n$, $F'(x)$ is a strictly diagonally dominant matrix, then F is a strictly diagonally dominant mapping.

Parts (a) and (d) are given by Moré and Rheinboldt [1970], and (e) is due to Moré [1970]. It may be noted that in some of these implications F-differentiability may be reduced to G-differentiability. Parts (b) and (c) represent results of Gale and Nikaido [1965] phrased in our terminology. It may be noted that (c) follows from (b) and Theorem 3.1(g). In fact, since $\partial_j f_i(x) \leq 0$ for $i \neq j$ and any $x \in R^n$, the mean value theorem applied to ψ_{ij} ensures that F is off-diagonally antitone.

As a typical application of Theorem 3.6, consider the two-point boundary value problem

$$u'' = \phi(t,u,u'),\ 0 < t < 1,\ u(0) = \alpha,\ u(1) = \beta \tag{3.2}$$

where ϕ is F-differentiable on $S = \{(t,u,p)^T \in R^3 \mid 0 \leq t \leq 1,\ u,p \in R^1\}$, and

$$\partial_2\phi(t,u,p) \geq 0,\ |\partial_3\phi(t,u,p)| \leq \gamma\ ,\quad \forall (t,u,p)^T \in S.$$

A simple discrete analog of (3.2) has the form

$$F: R^n \to R^n,\ Fx = Ax + h^2\Phi x + b$$

where $h = (n+1)^{-1}$, $tj = jh$, $j = 0,1,\ldots,n+1$,

$$A = \begin{pmatrix} 2 & -1 & & 0 \\ -1 & 2 & \ddots & \\ & \ddots & \ddots & -1 \\ 0 & & -1 & 2 \end{pmatrix},\ b = (-\alpha,0,\ldots,0,-\beta)^T,$$

and $\Phi:R^n \to R^n$ has the components $\phi_i(t_i, x_i, (2h)^{-1}(x_{i+1}-x_{i-1}))$, $i=1,\ldots,n$, with $x_0 = \alpha$, $x_{n+1} = \beta$. For $h < 2/\gamma$ it is readily verified that $F'(x) = A + h^2\Phi'(x)$ is always an M-matrix. In fact, $F'(x)$ is tridiagonal with positive diagonal and strictly negative first subdiagonals, and we have irreducible diagonal dominance. (See, e.g., Varga [1962].) Thus, F is an M-function. It turns out that F is also surjective; this can be shown in various ways; a simple proof follows from Theorem 4.7.

Since a linear mapping $A:R^n \to R^n$ is inverse isotone if and only if $A^{-1} \geq 0$, Theorem 3.6 suggests the conjecture that when $F:R^n \to R^n$ has, for any $x \in R^n$, a nonsingular F-derivative for which $F'(x)^{-1} \geq 0$, then F is inverse isotone. So far, this still represents an open problem, but there are several partial answers. We conclude this section with one of these; a second one is contained in the next section.

Theorem 3.7. Let $F:R^n \to R^n$ be convex and G-differentiable on R^n. Then F is inverse isotone if and only if, for any $x \in R^n$, $F'(x)$ is nonsingular and $F'(x)^{-1} \geq 0$.

The proof is given by Moré [1970].

4. Surjectivity

In this section we turn to the question when certain of the functions considered so far are surjective, that is, when the corresponding equation $Fx = z$ is solvable for any $z \in R^n$. The basic tool for our discussion will be the following "norm-coerciveness" theorem which appears to be due to Cacciopoli [1932] and which is also a special case of a more general result of Rheinboldt [1969a].

Theorem 4.1. Let $F:R^n \to R^n$ be a local homeomorphism. Then F is bijective if and only if F is norm-coercive in the sense that

$$\lim_{\|x\|\to\infty} \|Fx\| = +\infty. \tag{4.1}$$

In addition, we shall frequently use the well-known domain invariance theorem which ensures that a continuous, injective mapping $F:R^n \to R^n$ has an open range FR^n and is a homeomorphism from R^n onto FR^n.

As a direct application of Theorem 4.1, we prove the following generalization of a result of Sandberg and Willson [1969a]. Following Ortega and Rheinboldt [1970b], a mapping $\Phi:R^n \to R^n$ is diagonal if the ith component ϕ_i of Φ is a function of only the ith variable x_i, or, in other words, if the link-set Λ_Φ of the associated network of Φ is empty.

Theorem 4.2. Let $F:R^n \to R^n$ be a continuous P_0-function such that, independent of x, the off-diagonal link-functions are Lipschitz continuous; that is

$$|f_i(x+se^j)-f_i(x+te^j)| \leq \gamma_{ij}|s-t|, \qquad s,t \in R^1,\ x \in R^n,\ i \neq j. \tag{4.2}$$

Then $\hat{F} = F + \Phi$ is a surjective P-function for any diagonal, strictly isotone, and surjective mapping $\Phi:R^n \to R^n$, ϕ_i, $i = 1,\dots,n$.

Proof: From the definitions it follows readily that $\hat{F}$ is a P-function and, hence, injective. Moreover, each component of Φ is necessarily continuous on R^1 and thus also $\hat{F}$ is continuous. In order to apply Theorem 4.1, it remains to show only that $\hat{F}$ is norm-coercive. For this we proceed by induction with respect to the dimension n.

For $n = 1$, F is isotone and the statement is trivial. Assume therefore that the theorem is valid for dimension $n - 1$, and that $\{x^k\} \subset R^n$ is any sequence such that $\{\hat{F}x^k\}$ is bounded. By the definition of P_0-functions, there exists for any $k \geq 0$ an index $i_k \in N$ such that

$$x_{i_k}^k (f_{i_k}(x^k) - f_{i_k}(0)) \geq 0, \; x_{i_k}^k \neq 0,$$

and hence that

$$(4.3) \qquad x_{i_k}^k (\hat{f}_{i_k}(x^k) - \hat{f}_{i_k}(0)) = x_{i_k}^k (f_{i_k}(x^k) - f_{i_k}(0)) + x_{i_k}^k (\phi_{i_k}(x_{i_k}^k) - \phi_{i_k}(0))$$
$$\geq x_{i_k}^k (\phi_{i_k}(x_{i_k}^k) - \phi_{i_k}(0)).$$

We can select a subsequence of $\{x^k\}$ -- again denoted by $\{x^k\}$ -- such that i_k is constant, and, for ease of notation, that $i_k = n$ for all $k \geq 0$. Then (4.3) assumes the form

$$x_n^k \phi_n(x_n^k) \leq x_n^k (\hat{f}_n(x^k) - f_n(0)).$$

Since $\{\hat{F}x^k\}$ is bounded, and, say, $|\hat{f}_n(x^k) - f_n(0)| \leq c$, $k \geq 0$, it follows for $x_n^k \geq 0$ that $\phi_n(0) \leq \phi_n(x_n^k) \leq c$ or $0 \leq x_n^k \leq \phi_n^{-1}(c)$, while for $x_n^k < 0$ we obtain $\phi_n(0) \geq \phi_n(x_n^k) \geq -c$, and thus $0 \geq x_n^k \geq \phi_n^{-1}(-c)$. Altogether, therefore, $\{x_n^k\}$ is bounded.

Now consider the subfunction $\hat{G}: R^{n-1} \to R^{n-1}$ of $\hat{F}$ with the components

$$\hat{g}_i(x_1, \ldots, x_{n-1}) = \hat{f}_i(x_1, \ldots, x_{n-1}, 0) = f_i(x_1, \ldots, x_{n-1}, 0) + \phi_i(x_i),$$
$$i = 1, \ldots, n-1.$$

By Theorem 3.4, $\hat{G}$ satisfies again the conditions of the theorem, and it follows from the boundedness of $\{x_n^k\}$ and $\{\hat{F}x^k\}$ that

$$|\hat{g}_i(x_1^k,\ldots,x_{n-1}^k,0)|$$

$$\leq |\hat{f}_i(x^k)| + |f_i(x_1^k,\ldots,x_{n-1}^k,0)-f_i(x_1^k,\ldots,x_n^k)|$$

$$\leq |\hat{f}_i(x^k)| + \gamma_{in}|x_n^k| \leq \text{constant},\ i = 1,\ldots,n,\ k \geq 0.$$

Therefore, by induction hypothesis, $\{x_i^k\}$, $i = 1,\ldots,n-1$, must be bounded sequences, and this implies that $\hat{F}$ is indeed norm-coercive.

Note that the condition (4.2) certainly holds if F itself is uniformly Lipschitz-continuous. Thus the theorem applies, in particular, to the case $\hat{F} = A + \Phi$, where A is a P_0-matrix. This represents exactly the mentioned result of Sandberg and Willson [1969a]. At the same time, the one-dimensional example $Fx = e^x + x$ shows that (4.2) does not require F to be uniformly Lipschitzian.

In the case of inverse isotone, or M-functions, the rather stringent norm-coercivity assumption (4.1) can be replaced by the following condition of <u>order-coercivity</u>:

$$(4.4) \qquad \lim_{k\to\infty} \|Fx^k\| = +\infty \text{ whenever } \begin{cases} \lim\limits_{k\to\infty} \|x^k\| = +\infty \\ \text{and either } x^k \leq x^{k+1} \text{ or } x^k \geq x^{k+1} \\ \qquad \text{for all } k \geq 0. \end{cases}$$

For M-functions this was proved by Rheinboldt [1969b]. For the proof of the corresponding more general result on inverse isotone mappings the next simple observation will be useful:

Lemma 4.3. A continuous, inverse isotone mapping $F: R^n \to R^n$ is surjective if and only if

(4.5) $$\{y \in R^n \mid y = z + tv, -\infty < t < +\infty\} \subset FR^n$$

for some $v > 0$ and $z \in R^n$.

Proof: The necessity of the condition is trivial. For the proof of the sufficiency, observe first that F is a homeomorphism between R^n and FR^n and hence, by Theorem 4.1, that $FR^n = R^n$ if F is norm-coercive. Let $\{x^k\} \subset R^n$ be any sequence such that $\{Fx^k\}$ is bounded. Then, because of $v > 0$, we can choose constants α, β such that $\alpha v \leq Fx^k - z \leq \beta v$ for all $k \geq 0$, and hence, by (4.5), that

$$Fa = z + \alpha v \leq Fx^k \leq z + \beta v = Fb, \quad k = 0,1,\ldots$$

for certain $a,b \in R^n$. Therefore, it follows from the inverse isotonicity of F that $a \leq x^k \leq b$ for all k, which means that F is indeed norm-coercive.

The mentioned order-coercivity result has now the form:

Theorem 4.4. A continuous, inverse isotone mapping $F: R^n \to R^n$ is surjective if and only if it is order-coercive; that is, if and only if (4.4) holds.

Proof: If F is surjective, then Theorem 4.1 ensures that F is norm-coercive and hence order-coercive.

Conversely, let F be order-coercive and, with any fixed $z = Fx^0 \in FR^n$ and $\alpha > 0$, set $q(t): [0,1] \to R^n$, $q(t) = z + t\alpha e$, $0 \leq t \leq 1$, where $e = (1,1,\ldots,1)^T$. Since F is a homeomorphism between R^n and FR^n, we then have

$$\hat{t} = \sup \{t \in [0,1] \mid z + s\alpha e \in FR^n,\ s \in [0,t]\} > 0,$$

and $p(t) = F^{-1}q(t)$ is well-defined for $t \in [0,\hat{t}]$. Suppose that $\hat{t} < 1$ and let $\{t_k\} \subset R^1$ be such that $0 \leq t_k < t_{k+1} < \hat{t}$, $k = 0,1,\ldots$, and $\lim_{k\to\infty} t_k = \hat{t}$. Then $z \leq q(t_k) \leq q(t_{k+1}) \leq z + \hat{t}\alpha e$, $k = 0,1,\ldots$, and, by the inverse isotonicity of F, $p(t_k) \leq p(t_{k+1})$, $k = 0,1,\ldots$. But then the order-coercivity implies that $\{p(t_k)\}$ must be bounded, and, therefore, that $\lim_{k\to\infty} p(t_k) = \hat{x}$ exists. Now, by continuity, $F\hat{x} = q(\hat{t})$ and, since FR^n is open, $\hat{t}$ is clearly not maximal against assumption. Thus, necessarily, $\hat{t} = 1$ and, because $\alpha > 0$ was arbitrary, $z + te \in FR^n$ for all $t \geq 0$. Similarly it follows that $z + te \in FR^n$ for all $t \leq 0$ and now the result is a direct consequence of Lemma 4.3.

At the end of the previous section we mentioned the conjecture that when $F:R^n \to R^n$ is F-differentiable and $F'(x)^{-1} \geq 0$ for any $x \in R^n$, then F is inverse isotone. A partial answer to this question was given by Theorem 3.7. In the case of surjective mappings another partial result can be obtained with the help of a proof technique similar to that of the previous theorem.

Theorem 4.5. Suppose that $F:R^n \to R^n$ is continuously F-differentiable and that, for any $x \in R^n$, $F'(x)$ is nonsingular and satisfies $F'(x)^{-1} \geq 0$. Then F is inverse isotone and surjective if and only if it is order-coercive.

Proof: The necessity part was proved in Theorem 4.4. Suppose therefore that F is order-coercive. By the inverse-function theorem F is a local homeomorphism. More specifically, for any $x^0 \in R^n$ there exist open neighborhoods U of x^0 and V of Fx^0 such that the restriction F_U of F

to U is a homeomorphism from U onto V and that the inverse $G = F_U^{-1}: V \to U$ is again continuously F-differentiable with $G'(y) = F'(Gy)^{-1} \geq 0$ for any $y \in V$. But then, for any $y \leq z$, $y,z \in V$, it follows from

$$Gz - Gy = \int_0^1 G'(y+t(z-y))(z-y)dt \geq 0$$

that G is isotone on V, and hence, by Theorem 3.2, that F_U is inverse isotone on U. In other words, for any $x^0 \in R^n$ there exists an open neighborhood U of x^0 in which F is inverse isotone.

Let now $y^0 = Fx^0$, $u \geq 0$, $u \neq 0$, be any vectors, and set $q:[0,1] \to R^n$, $q(t) = y^0 + tu$, $0 \leq t \leq 1$. By the local homeomorphism property, there is a $t_1 > 0$ and a continuous function $p:[0,t_1] \to R^1$ such that $p(0) = x^0$, and $Fp(t) = q(t)$, $t \in [0,t_1]$. If $t_1 < 1$, we can repeat this argument and continue p beyond t_1 to some $t_2 > t_1$, etc. This continuation process ensures the existence of a continuous $p:[0,\hat{t}) \to R^1$ such that $p(0) = x^0$, and $Fp(t) = q(t)$ for $t \in [0,\hat{t})$. Let $\hat{t} \in (0,1]$ be the maximal value up to which p can be extended. For any $r,s \in [0,\hat{t})$, $r < s$, the set $p([r,s])$ is compact and hence can be covered by finitely many open sets $U_1,\ldots,U_m$ in each of which F is inverse isotone. More specifically, we can select points $r = r_1 < r_2 < \ldots < r_{n+1} = s$, such that $\{p(t) \mid r_i \leq t \leq r_{i+1}\} \subset U_i$, $i = 1,\ldots,m$. Then $q(r_{i+1}) \geq q(r_i)$ and the inverse isotonicity of F in each U_i imply that $p(r_{i+1}) \geq p(r_i)$, $i = 1,\ldots,m$, and hence that

$$p(s) - p(r) = \sum_{i=1}^{m} (p(r_{i+1})-p(r_i)) \geq 0, \quad 0 \leq r < s < \hat{t}. \tag{4.6}$$

Suppose now that $\hat{t} < 1$ and let $\{t_k\} \subset R^1$ be such that $0 \leq t_k < t_{k+1} < \hat{t}$, $k = 0,1,\ldots$, and $\lim_{k\to\infty} t_k = \hat{t}$. By (4.6) we have $p(t_k) \leq p(t_{k+1})$, $k = 0,1,\ldots$, and hence $y^0 \leq q(t_k) \leq y^0 + \hat{t}u$, $k \geq 0$, together with the order-coercivity, implies that $\{p(t_k)\}$ is bounded and, therefore, that $\lim_{k\to\infty} p(t_k) = \hat{x}$ exists. Because of the continuity of F, we now have $F\hat{x} = q(t)$ and the openness of FR^n contradicts the maximality of $\hat{t}$.

With this we have shown that $y^0 + u \in FR^n$ for any $y^0 \in FR^n$ and $u \geq 0$, $u \neq 0$. By the same argument it follows that also $y^0 - v \in FR^n$ whenever $y^0 \in FR^n$ and $v \geq 0$, $v \neq 0$. Since any point $y \in R^n$ can be written in the form $y = y^0 + u - v$, where $y^0 \in FR^n$ and $u \geq 0$, $v \geq 0$, we see that $y \in FR^n$ and hence that $FR^n = R^n$. Therefore, by Theorem 4.1, F is bijective and hence a homeomorphism from R^n onto itself. Thus, if $Fy \geq Fx$, then either $y = x$ or $u = Fy - Fx \geq 0$, $u \neq 0$. With $p(t) = F^{-1}(x+tu)$, $0 \leq t \leq 1$, the argument leading to (4.6) shows that $y = p(1) \geq p(0) = x$, and, hence, that F is inverse isotone.

For continuous M-functions $F:R^n \to R^n$, Rheinboldt [1969b] has proved two surjectivity results which are based on Theorem 4.3; in other words, the assumptions placed upon F are sufficient to prove the order coercivity. For later reference we quote here without proof one of these results, and more specifically the one which represents a continuation of Theorem 3.5.

Theorem 4.6. Let $F:R^n \to R^n$ be continuous and off-diagonally antitone, and suppose that for some continuous, strictly isotone, surjective functions $h_j:R^1 \to R^1$, $j = 1,\ldots,n$, the mapping $P:R^n \to R^n$ with components $p_i(t) = f_i(h_1(x_1+t),\ldots,h_n(x_n+t))$, $i = 1,\ldots,n$, is isotone for any fixed $x \in R^n$. Assume further that for every node i of Ω_F there is a path (3.1) to some node ℓ such that, for any $x \in R^n$, the link functions $\psi_{i_j i_{j+1}}$, $j = 0,\ldots,m$, as well as the component p_ℓ are strictly isotone and surjective. Then F is a surjective M-function.

As a simple application to differentiable mappings, we present the following corollary:

Theorem 4.7. Let $F:R^n \to R^n$ be off-diagonally antitone and F-differentiable, and suppose that $F'(x)u \geq v \geq 0$ for all $x \in R^n$ and some fixed $u > 0$, $v \geq 0$. Assume further that for any node i of Ω_F there is a path (3.1) to some node ℓ such that $v_\ell > 0$ and

$$(4.7) \qquad \partial_{i_{j+1}} f_{i_j}(x) \leq a_j < 0, \; j = 0,\ldots,m,$$

with constant $a_0,\ldots,a_m$. Then F is a surjective M-function.

With $h_j(t) = t/u_j$, $j = 1,\ldots,n$, this theorem reduces immediately to the previous one.

As mentioned earlier, Theorem 3.7 provides an easy means of verifying that the discrete analog of (3.2)--as considered in Section 3--is a surjective M-function.

Clearly, there are various other corollaries of Theorem 4.6 along the lines of the previous result. Rather than to detail these possibilities, we end this section with a somewhat different observation. The discussion in this section was based on the norm-coerciveness theorem 4.1. As mentioned, Rheinboldt [1969a] obtained this theorem as a special case of a more general continuation theory; another special case of this theory is the well-known Hadamard theorem which states that a continuously F-differentiable mapping $F:R^n \to R^n$ is surjective if $F'(x)$ is nonsingular and $\|F'(x)^{-1}\| \leq \gamma$ for all $x \in R^n$. The rather simple and direct proof of the next theorem shows that also the Hadamard theorem can be used to obtain surjectivity results of the type considered here.

Theorem 4.8. Let $F:R^n \to R^n$ be off-diagonally antitone and continuously F-differentiable. Suppose further that for any $x \in R^n$ there is a vector $u(x) > 0$ such that $u(x) \leqslant u$ and $F'(x)u(x) \geqslant v > 0$ with fixed $u > 0$, $v > 0$. Then F is a surjective M-function.

Proof: By the earlier-cited characterization result for M-matrices, $F'(x)$ is an M-matrix for any $x \in R^n$, and hence Theorem 3.6 ensures that F is an M-function. Moreover, $F'(x)^{-1} \geqslant 0$ implies that $0 \leqslant F'(x)^{-1}v \leqslant u$. Let $F'(x)^{-1} = (b_{ij}(x))$, then

$$0 \leqslant b_{ik}(x)v_k \leqslant \sum_{j=1}^{n} b_{ij}(x)v_j \leqslant u_i, \quad i,k = 1,\ldots,n,$$

and $v_k > 0$ shows that

$$\|F'(x)^{-1}\|_\infty \leqslant n(\max_i u_i)/(\min_k v_k),$$

and, therefore, that the Hadamard theorem applies.

5. Boundary Value Problems

For a mapping $F:R^n \to R^n$ representing an equilibrium flow on a given network, the following problem is basic: A state vector x is to be determined which satisfies certain specified conditions at the boundary nodes of the network and for which the efflux from all other nodes equals a prescribed value. In line with Birkhoff and Kellogg [1966], we consider here the case when the state at the boundary nodes is a given function of the efflux from that node. The system of equations to be solved has then the form

$$(5.1)\qquad \begin{cases} x_i - h_i(f_i(x)) = z_i, & i \in N_b \\ f_i(x) = z_i, & i \notin N_b \end{cases}$$

where $N_b \subset N$ is the set of boundary nodes. It is no restriction to assume always that $N_b = \{1,\ldots,m\}$, $1 \leq m \leq n$.

Any system of the form (5.1) is a (specific) boundary value problem for F with respect to the boundary set N_b. We shall combine the given functions h_i into a mapping $H:R^1 \to R^m$ with components $h_1,\ldots,h_m$, and denote (5.1) by $\{H,z\}$ where $z \in R^n$. Correspondingly, $x = \text{sol}\,\{H,z\}$ designates any solution of (5.1).

The simplest type of boundary value problem is obtained when $H \equiv 0$, that is, when the boundary conditions reduce to $x_i = z_i$, $i \in N_b$; this will be called the Dirichlet boundary value problem. In this case, the system (5.1) is equivalent with

$$(5.2)\qquad f_i(z_1,\ldots,z_m,x_{m+1},\ldots,x_n) = z_i,\ i = m+1,\ldots,n,$$

and hence we are again led to consider the properties of the subfunctions of F.

In many instances it is of interest to provide results for all boundary value problems $\{H,z\}$ obtained by letting $z \in R^n$ be any vector and H any function from a given class $\mathcal{H}$ of mappings. For abbreviation, we denote the collection of all problems $\{H,z\}$ of this type by $B(F,N_b,\mathcal{H})$, or $B(\mathcal{H})$ for short, if F and N_b are fixed. The two basic problems connected with such a class $B(\mathcal{H})$ are, of course, as in the case of differential equations, the existence and uniqueness of solutions for any $\{H,z\} \in B(\mathcal{H})$.

In view of the above indicated connection between the Dirichlet problems and the subfunctions of the mapping F, the following result is a direct consequence of Theorem 3.4:

Theorem 5.1. Let $F:R^n \to R^n$ be a P-, or M-function. Then the solution of any Dirichlet boundary value problem $\{0,z\}$, $z \in R^n$ of F is unique provided it exists.

For continuous, surjective M-functions $F:R^n \to R^n$, Rheinboldt [1969b] has shown that also any subfunction is again a surjective M-function. Thus, in that case any Dirichlet problem of F always has a unique solution. The proof of this result was based on a characterization of surjective M-functions in terms of the convergence of the Jacobi process. We give here a simple, direct proof based on the order coercivity theorem 4.3 and on the fact that the subfunctions of an M-function are again M-functions.

Theorem 5.2. Let $F:R^n \to R^n$ be a continuous surjective M-function. Then also any subfunction of F is again a surjective M-function.

Proof: We prove the result for the subfunction

$$G:R^{n-1} \to R^{n-1},\ g_i(x_1,\ldots,x_{n-1}) = f_i(x_1,\ldots,x_{n-1},c_n),\ i=1,\ldots,n-1;$$

the general case then follows by repeated application of this result and by appropriate permutations of the variables and components. Because of Theorems 3.4 and 4.3 we need to show only that G is order-coercive. Let $\bar{x}^k = (x_1^k,\ldots,x_{n-1}^k)^T \in R^{n-1}$ be any monotonically increasing sequence such that $\{G\bar{x}^k\}$ is bounded. If, say, $b_i \geq g_i(\bar{x}^k)$, $i = 1,\ldots,n-1$, then

$$b_i \geq f_i(x_1^k,\dots,x_{n-1}^k,c_n),\ i = 1,\dots,n-1,$$
$$b_n = f_n(x_1^0,\dots,x_{n-1}^0,c_n) \geq f_n(x_1^k,\dots,x_{n-1}^k,c_n), \qquad k = 0,1,\dots,$$

and hence it follows, with $v = F^{-1}(b_1,\dots,b_n)$, that $v_i \geq x_i^k$, $i = 1,\dots,n-1$, $k = 0,1,\dots$. Therefore, $\{\bar{x}^k\}$ is bounded, and the same result can be obtained if $\{\bar{x}^k\}$ is monotonically decreasing. This shows that G is order coercive and thus surjective.

It may be interesting to note that the subfunctions of surjective, inverse isotone mappings need neither be inverse isotone nor surjective. In fact, it is easily verified that

$$F:R^3 \to R^3,\ Fx = \begin{pmatrix} -x_1 + x_2 \\ x_1 - x_2 + x_3 \\ x_1 \qquad - x_3^3 \end{pmatrix}$$

is inverse isotone and surjective, while the subfunction $G_1:R^1 \to R^1$, $g_1(x_1) = -x_1$ is not inverse isotone and the subfunction

$$G_2:R^2 \to R^2,\ G_2x = \begin{pmatrix} -x_1 + x_2 \\ x_1 - x_2 \end{pmatrix}$$

is not surjective.

In the case of elliptic partial differential equations, the validity of a maximum principle is an important tool in the study of the corresponding boundary value problems. In line with Rheinboldt [1969b]--where a somewhat different terminology was used--we define a maximum principle for network boundary value problems as follows:

Definition 5.3. Consider a class $B(F,N_b,\mathcal{H})$ of boundary value problems for a mapping $F:R^n \to R^n$ with respect to some set of boundary nodes $N_b \subset N$. The class $B(\mathcal{H})$ admits the maximum principle if for any $\{H^i,z^i\} \varepsilon B(\mathcal{H})$, $i=1,2$, with $H^1(t) \leqslant H^2(t)$, $t \varepsilon R^1$, and $z^1 \leqslant z^2$, it follows that $x^1 \leqslant x^2$ for any $x^i = \text{sol}\,\{H^i,z^i\}$, $i = 1,2$.

For continuous, off-diagonally antitone F and the class $\mathcal{H} = \mathcal{A}$ of all continuous, antitone $H:R^1 \to R^m$, Rheinboldt [1969b] proved a theorem ensuring the validity of the maximum principle. This result was based on the observation that $B(F,N_b,\mathcal{A})$ admits the maximum principle if and only if, for any $H \varepsilon \mathcal{A}$, the mapping

$$(5.3) \qquad F^H:R^n \to R^n,\ f_i^H(x) = \begin{cases} x_i - h_i(f_i(x)), & i = 1,\dots,m \\ f_i(x) & , i = m+1,\dots,n, \end{cases}$$

is inverse isotone.

In the differentiable case there is a simpler version of this result for which a proof can be obtained directly from Theorem 4.7. We denote by $\mathcal{A}^*$ the class of all F-differentiable, antitone mappings $H:R^1 \to R^m$.

Theorem 5.4. Let $F:R^n \to R^n$ be off-diagonally antitone and F-differentiable and suppose that $F'(x)u \geqslant v \geqslant 0$ for any $x \varepsilon R^n$ and fixed $u > 0$, $v \geqslant 0$. Assume further that for any $i \notin N_b = \{1,\dots,m\}$ there exists a path (3.1) from i to a boundary node ℓ for which (4.7) holds. Then every boundary value problem $\{H,z\} \varepsilon B(F,N_b,\mathcal{A}^*)$ has a unique solution and the class $B(\mathcal{A}^*)$ admits the maximum principle.

To apply Theorem 4.7 we need to note only that for any $H \in \mathcal{A}^*$ the function F^H of (5.3) is off-diagonally antitone and F-differentiable, and that $(F^H)'(x)u \geq \hat{v} \geq 0$ with $\hat{v}_i = u_i > 0$ for $i \in N_b$ and $\hat{v}_i = v_i$ for $i \notin N_b$.

6. Iterative Processes

This section is not intended to give a survey about the iterative solution of the equation $Fx = z$ when F belongs to any one of the function classes discussed here. There are many relevant results in the literature, and such a survey would, by necessity, be rather extensive (see, e.g., Ortega and Rheinboldt [1970b]). Even if we restrict ourselves to global convergence theorems, the list of results still remains surprisingly lengthy. For the Jacobi- or Gauss-Seidel processes it would include a well-known theorem of Schechter [1962], which in our terminology concerns a special type of P-function, new results of Moré [1970] for diagonally dominant functions, and also the result of Rheinboldt [1969b] for surjective M-functions, cited in the introduction. In addition, we would have to mention the global convergence theorem of Greenspan and Parter [1965] for the Newton-one-step Gauss-Seidel process, which applies to a simple type of M-function, and also the Newton-convergence theorem of Baluev [1952] which, as we shall see, assumes the inverse isotonicity of F.

Instead of going into further details about these and other related results, we shall consider here a particular class of implicit iterative processes of the form

(6.1) $$G(x^{k+1},x^k) = z,\ k = 0,1,\ldots,$$

which represents a nonlinear generalization of the family of linear methods obtained from regular splittings.

For the iterative solution of a linear system $Ax = z$, Varga [1962] introduced a regular splitting of the matrix A as a decomposition $A = B - C$ in which B is nonsingular and $B^{-1} \geq 0$ as well as $C \geq 0$. He then showed that for such splittings the iterative process $x^{k+1} = B^{-1}Cx^k + B^{-1}z$, $k = 0,1,\ldots,$ converges (for any $x^0 \in R^n$) to the solution of $Ax = z$ if $A^{-1} \geq 0$. This covers, in particular, the well-known result about the convergence of the Jacobi and the Gauss-Seidel process when A is an M-matrix.

The mentioned similarity between the behavior of the M-matrices and M-functions suggests the idea of generalizing these regular splittings to nonlinear mappings. In direct analogy to the linear definition we introduce this generalization as follows:

Definition 6.1. A mapping $G:D_0 \times D_0 \subset R^n \times R^n \to R^n$ is a regular iteration function for $F:D \subset R^n \to R^n$ on the subset D_0 of D if

(6.2a) $G(x,x) = Fx$, for any $x \in D_0$,

(6.2b) $G(\cdot,x):D_0 \to R^n$ is inverse isotone, for any fixed $x \in D_0$,

(6.2c) $G(y,\cdot):D_0 \to R^n$ is antitone, for any fixed $y \in D_0$.

Note that in the linear case this definition reduces exactly to that of a regular splitting. Note also that in the nonlinear case it is evidently necessary to add some assumption about the solvability of the equation

$G(y,x) = z$, for given $x \in D_0$, before we can hope to establish general convergence results about the (implicit) iterative process (6.1).

In the remainder of this section we shall denote order intervals in R^n by $\langle u,v\rangle = \{x \in R^n \mid u \leq x \leq v\}$, and we write $x^k \uparrow x^*$, $k \to \infty$, as an abbreviation for $x^k \leq x^{k+1}$, $k = 0,1,\ldots$, $\lim_{k\to\infty} x^k = x^*$, and, similarly, $x^k \downarrow x^*$, $k \to \infty$, for $x^k \geq x^{k+1}$, $k = 0,1,\ldots$, $\lim_{k\to\infty} x^k = x^*$. Moreover, we will use the following well-known result of Kantorovich [1939]:

Lemma 6.2. Let $H:\langle x^0,y^0\rangle \subset R^n \to R^n$ be continuous and isotone, and assume that $x^0 \leq Hx^0$ and $y^0 \geq Hy^0$. Then the sequences $x^{k+1} = Hx^k$, $y^{k+1} = Hy^k$, $k = 0,1,\ldots$, satisfy $x^k \uparrow x^*$, $k \to \infty$, and $y^k \downarrow y^*$, $k \to \infty$, where $x^0 \leq x^* = Hx^* \leq y^* = Hy^* \leq y^0$.

The next theorem extends to processes of the form (6.1) a corresponding monotone convergence result for Jacobi and Gauss-Seidel processes proved by Rheinboldt [1969b].

Theorem 6.3. Given $F:D \subset R^n \to R^n$, suppose that for some $z \in R^n$ there are $x^0,y^0 \in D$ such that $x^0 \leq y^0$, $J = \langle x^0,y^0\rangle \subset D$ and $Fx^0 \leq z \leq Fy^0$. Let $G:J \times J \to R^n$ be a continuous, regular iteration function for F on J with the property that $z \in G(J;x)$ for any fixed $x \in J$. Then the sequences $\{y^k\}$ and $\{x^k\}$ specified by (6.1) and starting from y^0 and x^0, respectively, are well-defined and satisfy $x^k \uparrow x^*$, $k \to \infty$, $y^k \downarrow y^*$, $k \to \infty$, where $x^0 \leq x^* \leq y^* \leq y^0$ and x^*,y^* are both solutions of $Fx = z$.

Proof: By assumption $G(\cdot,x) = z$ has a solution $y \in J$ and, because of (6.2b), this solution is unique. Hence the mapping $H:J \to J$ satisfying $G(Hx,x) = z$ is well-defined and clearly continuous. Moreover, by (6.2c) it follows from $x^0 \leq x \leq y \leq y^0$ that

$$G(Hy,y) = z = G(Hx,x) \geq G(Hx,y)$$

and hence, by (6.2b), that $Hy \geq Hx$. Therefore, H is isotone on J. Finally, again by (6.2b), $G(x^0,x^0) = Fx^0 \leq z = G(Hx^0,x^0)$ implies that $x^0 \leq Hx^0$, and, similarly, we obtain from $G(y^0,y^0) = Fy^0 \geq z = G(Hy^0,y^0)$ that $y^0 \geq Hy^0$. The convergence statement is now a direct consequence of Lemma 6.2, and $x^* = Hx^*$ and $y^* = Hy^*$ are equivalent with $Fx^* = G(x^*,x^*) = z$ and $Fy^* = G(y^*,y^*) = z$.

As a corollary we obtain the following global convergence result.

Theorem 6.4. Let $F:R^n \to R^n$ be continuous, inverse isotone, and surjective. Suppose, further, that $G:R^n \times R^n \to R^n$ is a regular iteration function for F on R^n with the property that $G(\cdot,x):R^n \to R^n$ is surjective for any fixed $x \in R^n$. Then, for any $z \in R^n$ and any initial point $x^0 \in R^n$, the process (6.1) converges to the unique solution $x^* \in R^n$ of $Fx = z$.

Proof: Note first that (6.2b) and (6.2c) together with the surjectivity of $G(\cdot,x)$ imply the existence of a continuous, isotone mapping $H:R^n \to R^n$ for which $G(Hx,x) = z$ for all $x \in R^n$. Moreover, for any $x^0 \in R^n$ the sequence $\{x^k\} \subset R^n$ given by (6.1) is uniquely defined by $x^{k+1} = Hx^k$, $k = 0,1,\ldots$.

Let $a,b \in R^n$ be the vectors with the components

$$a_j = \min\ (f_j(x^0), z_j),\ b_j = \max\ (f_j(x^0), z_j),\ j=1,\ldots,n,$$

and set $u^0 = F^{-1}a$, $v^0 = F^{-1}b$. Then $Fu^0 \leqslant z \leqslant Fv^0$ as well as $Fu^0 \leqslant x^* \leqslant Fv^0$, and it follows from the inverse isotonicity of F that $u^0 \leqslant x^* \leqslant v^0$, as well as $u^0 \leqslant x^0 \leqslant v^0$. Consider now the sequences $\{u^k\}$, $\{x^k\}$, $\{v^k\}$ given by (6.1) and starting from u^0, x^0, and v^0, respectively. Then $u^{k+1} = Hu^k$, $x^{k+1} = Hx^k$, $v^{k+1} = Hv^k$, $k = 0,1,\ldots$, and, because of the isotonicity of H, we obtain from $u^0 \leqslant x^0 \leqslant v^0$, by induction, that $u^k \leqslant x^k \leqslant v^k$, $k = 0,1,\ldots$. Moreover, from Theorem 6.3 it follows that $u^k \uparrow u^*$, $k \to \infty$ and $v^k \downarrow v^*$, $k \to \infty$ where $u^0 \leqslant u^* \leqslant v^* \leqslant v^0$ and $Fu^* = z$, $Fv^* = z$. The injectivity of F then implies that $u^* = v^* = x^*$ and hence necessarily that, $\lim_{k\to\infty} x^k = x^*$.

This result covers as a special case the global convergence of the Jacobi and the Gauss-Seidel process for continuous surjective M-functions mentioned in the introduction. Instead of going into details of that case we extend it immediately to the corresponding block processes.

With $n_1 + n_2 + \ldots + n_p = n$, $n_j \geqslant 1$, $p \geqslant 1$, consider R^n as the product-space $R^{n_1} \times R^{n_2} \times \ldots \times R^{n_p}$ and let $P_i : R^n \to R^{n_i}$, $i = 1,\ldots,p$, denote the corresponding natural projections. Then any $x \in R^n$ may be partitioned in the form $x = (x^1,\ldots,x^p)$ where $x^i = P_i x$, $i = 1,\ldots,p$, and, similarly, we define the block-components $F^i : R^n \to R^{n_i}$ of any mapping $F : R^n \to R^n$ by $F^i x = P_i Fx$, $i = 1,\ldots,p$.

For the solution of the equation $Fx = z$, the block-Gauss-Seidel process, with respect to the particular partition, has now the form:

$$
(6.3)\quad \begin{cases} \text{Determine a solution } x^i \in R^{n_i} \text{ of} \\ F^i((x^1)^{k+1},\dots,(x^{i-1})^{k+1}, x^i, (x^{i+1})^k,\dots,(x^n)^k) = z^i \\ \text{and set } (x^i)^{k+1} = x^i,\ i = 1,\dots,p,\ k = 0,1,\dots \ . \end{cases}
$$

Analogously, the block Jacobi process can be defined, and the form of the relaxation-versions of both processes should be self-evident. For reasons of space, we restrict ourselves here to (6.3), especially since in the mentioned other cases the discussion remains essentially the same.

Observe now that with

$$
(6.4)\qquad G:R^n \times R^n \to R^n,\ P_iG(y,x) = F^i(y^1,\dots,y^i,x^{i+1},\dots,x^n),\ i=1,\dots,p,
$$

the process (6.3) assumes the general form (6.1). For M-functions F, the following result then ensures the applicability of Theorems 6.3 and 6.4 to the block Gauss-Seidel process (6.3).

Theorem 6.5. Let $F:R^n \to R^n$ be an M-function; then the mapping $G:R^n \times R^n \to R^n$ defined by (6.4) is a regular iteration function for F on R^n. If, in addition, F is continuous and surjective, then $G(\cdot,x):R^n \to R^n$ is surjective for any fixed $x \in R^n$.

Proof: The conditions (6.2a) and (6.2c) are evidently satisfied for G, and hence, in order to complete the proof of the first statement, it remains to show only that $G(\cdot,x):R^n \to R^n$ is inverse isotone for any $x \in R^n$. Let $G(v,x) \geq G(u,x)$ for some $x,u,v \in R^n$. By Theorem 3.4, the subfunction $F^1(\cdot,x^2,\dots,x^n):R^{n_1} \to R^{n_1}$ of F is again an M-function, and, hence,

$P_1G(v,x) \geq P_1G(u,x)$ implies that $v^1 \geq u^1$. To proceed by induction, suppose that $v^i \geq u^i$ for $i = 1,\ldots,k-1$ and some k with $2 \leq k \leq n$. Then the off-diagonal antitonicity of F ensures that

$$F^k(u^1,\ldots,u^{k-1},v^k,x^{k+1},\ldots,x^n) \geq F^k(v^1,\ldots,v^k,x^{k+1},\ldots,x^n)$$

$$= P_kG(v,x) \geq P_kG(u,x) = F^k(u^1,\ldots,u^k,x^{k+1},\ldots,x^n).$$

Since $F^k(u^1,\ldots,u^{k-1},\cdot,x^{k+1},\ldots,x^n):R^{n_k} \to R^{n_k}$ is again an M-function, this shows that also $v^k \geq u^k$, and hence altogether, that $v \geq u$.

For the proof of the second part, assume that F is a continuous, surjective M-function. Then, by Theorem 5.2, also any subfunction of F has the same properties. Hence, for given $x,z \in R^n$, the equation

$$P_1G(y,x) = F^1(y^1,x^2,\ldots,x^n) = z^1$$

has a unique solution $y^1 \in R^{n_1}$. If for some k, with $2 \leq k \leq n-1$, vectors $y^i \in R^{n_i}$, $i = 1,\ldots,k-1$, have already been found with the property that

$$(6.5) \qquad P_iG(y,x) = F^i(y^1,\ldots,y^i, x^{i+1},\ldots,x^n) = z^i, \; i = 1,\ldots,k-1,$$

then the surjectivity of $F^k(y^1,\ldots,y^{k-1},\cdot,x^{k+1},\ldots,x^n)$ ensures the existence of a $y^k \in R^{n_k}$ for which (6.5) holds with $i = k$. Hence, $G(y,x) = z$ has a unique solution $y \in R^n$ and the proof is complete.

Theorems 6.4 and 6.5 together state the global convergence of the block Gauss-Seidel process (6.3) for continuous, surjective M-functions. As mentioned above, this result also carries over to the block-Jacobi process and the underrelaxed versions of both these block methods, and the proofs for these cases are essentially the same as that for the Gauss-Seidel method.

These results raise the question whether some of the other global convergence theorems mentioned in the beginning of this section might also be subsumed under Theorem 6.4. This is not the case due to the fairly restrictive nature of the regular iteration functions. There appear to be many possible modifications of Definition 6.1, but it is doubtful whether any one of them actually covers a broader class of methods and not just again only a few specific convergence theorems. We end this section with some results about one such generalized form of the regular iteration functions. Since this discussion is primarily intended to be an example, no attempt was made to phrase these results in their most general form.

Theorem 6.6. Let $F:R^n \to R^n$ be inverse isotone and $G:R^n \times R^n \to R^n$ a continuous mapping with the properties

(6.6a) $G(x,x) = Fx$ for any $x \in R^n$

(6.6b) $G(\cdot,x):R^n \to R^n$ is inverse isotone and surjective for any $x \in R^n$

(6.6c) $G(x,x) \geq G(y,x)$, $x,y \in R^n$ implies that $G(y,y) \geq G(y,x)$.

If for some x^0, y^0, $z \in R^n$ we have $Fx^0 \leq z \leq Fy^0$, then the sequence $\{y^k\}$ given by (6.1) and starting from y^0 satisfies $y^k \downarrow x^*$, $k \to \infty$, where $x^* \in \langle x^0, y^0 \rangle$ is the (unique) solution of $Fx = z$ in R^n.

Proof: By (6.6b) the mapping $H:R^n \to R^n$ satisfying $G(Hx,x) = z$ is well-defined and certainly continuous. Then $\{y^k\}$ is uniquely specified by $y^{k+1} = Hy^k$, $k = 0,1,\ldots$. Moreover, from (6.6c) it follows that when

$G(x,x) = Fx \geq z = G(Hx,x)$, then also $FHx = G(Hx,Hx) \geq G(Hx,x) = z$. Thus we obtain from $Fy^0 \geq z$, by induction, that $Fy^k \geq z \geq Fx^0$, $k = 0,1,\ldots$. Now the inverse isotonicity of F implies that $y^k \geq x^0$, for all $k \geq 0$, which, in turn, ensures the existence of $\lim_{k\to\infty} y^k = x^* \in \langle x^0,y^0\rangle$. By the continuity of G we then have $Fx^* = G(x^*,x^*) = z$ and, clearly, x^* is unique.

Note that the analogous result holds for the lower sequence $\{x^k\}$ starting from x^0 provided that all inequalities in (6.6c) are reversed. Any regular iteration function satisfies (6.6c) as well as the corresponding reversed implication. In fact, if $G(x,x) \geq G(y,x)$, $x,y \in R^n$, then we obtain from (6.2b) that $x \geq y$ and hence from (6.2c) that $G(y,y) \geq G(y,x)$. The analogous argument applies when the inequalities are reversed.

As an application of this result we give the following theorem, proved in more generality by Ortega and Rheinboldt [1967].

<u>Theorem 6.7</u>. Let $F:R^n \to R^n$ be continuously differentiable, convex, and inverse isotone, and assume that $F'(x) = B(x) - C(x)$ is, for any $x \in R^n$, a regular splitting of $F'(x)$ with continuous $B:R^n \to L(R^n)$. If for given x^0,y^0, $z \in R^n$ we have $Fx^0 \leq z \leq Fy^0$ (and thus $x^0 \leq y^0$), then the sequence

$$(6.7) \qquad y^{k+1} = y^k - B(y^k)^{-1}(Fy^k-z),\ k = 0,1,\ldots$$

satisfies $y^k \downarrow x^*$, $k \to \infty$, where $x^* \in \langle x^0,y^0\rangle$ is the (unique) solution of $Fx = z$ in R^n.

In order to use Theorem 6.6, we note first that for

$$G:R^n \times R^n \to R^n,\ G(y,x) = B(x)(y-x) + Fx$$

the process (6.7) is equivalent with (6.1). Clearly G satisfies (6.6a) and (6.6b); for the proof of (6.6c) let $G(x,x) \geq G(y,x) = w$ for some $x,y \in R^n$. This inequality is equivalent with $0 \geq B(x)(y-x)$, and hence, $B(x)^{-1} \geq 0$ implies that $x \geq y$. With the help of the convexity inequality

$$(6.8) \qquad Fy - Fx \geq F'(x)(y-x), \; x,y \in R^n,$$

we find then that indeed

$$\begin{aligned} G(y,y) = Fy &\geq Fx + F'(x)(y-x) \\ &= Fx - (B(x)-C(x))B(x)^{-1}(Fx-w) \\ &= C(x)B(x)^{-1}(Fx-w) + w \geq w = G(y,x). \end{aligned}$$

Note that G is, in general, not a regular iteration function.

Ortega and Rheinboldt [1967] have shown that this result contains as a corollary the global convergence theorem for Newton's method of Baluev [1952] mentioned in the beginning of this section. It may be interesting that Theorem 3.7 allows us to state this theorem in a slightly modified, and yet equivalent, form:

Theorem 6.8. Let $F:R^n \to R^n$ be continuously differentiable, convex, and inverse isotone. If $Fx = z$ has a solution x^*, then, for any $x^0 \in R^n$, the Newton iterates $x^{k+1} = x^k - F'(x^k)^{-1}(Fx^k - z)$, $k = 0,1,\ldots,$ are well-defined and satisfy $x^k \geq x^{k+1}$, $k = 1,2,\ldots,$ and $\lim_{k\to\infty} x^k = x^*$.

By Theorem 3.7, $F'(x)$ is always nonsingular and $F'(x)^{-1} \geq 0$. Thus the mapping $G:R^n \times R^n \to R^n$, $G(y,x) = F'(x)(y-x) + Fx$ clearly satisfies (6.6a) and (6.6b), while (6.6c) is simply (6.8). From (6.8) also follows that $Fx^1 \geq Fx^0 + F'(x^0)(x^1-x^0) = z$, and hence, that Theorem 6.6 applies on the order interval $\langle x^*,x^1\rangle$.

References

Baluev, A. [1952], On the method of Chaplygin, Dokl. Akad. Nauk SSSR 83, 781-784.

Bers, L. [1953], On mildly nonlinear partial difference equations of elliptic type, J. Res. Nat. Bur. Stand. 51, 229-236.

Birkhoff, G., and Kellogg, B. [1966], Solution of equilibrium equations in thermal networks, Proc. Symp. General. Networks, Brooklyn Polyt. Press.

Cacciopoli, R. [1932], Sugli elementi uniti delle trasformazioni funzionali, Rend. Sem. Mat. Univ. Padova 3, 1-15.

Fan, Ky [1958], Topological proofs for certain theorems on matrices with nonnegative elements, Monatsk. Math. 62, 219-237.

Fiedler, M., and Ptak, V. [1962], On matrices with non-positive off-diagonal elements and positive principal minors, Czech. Math. J. 12, 382-400.

Fiedler, M., and Ptak, V. [1966], Some generalizations of positive definiteness and monotonicity, Num. Math. 9, 163-172.

Gale, D., and Nikaido, H. [1965], The Jacobian matrix and global univalence of mappings, Mathem. Annalen 159, 81-93.

Greenspan, D., and Parter, S. [1965], Mildly nonlinear elliptic partial differential equations and their numerical solution II, Num. Math. 7, 129-147.

Kantorovich, L. [1939], The method of successive approximations for functional equations, Acta Math. 71, 63-97.

Karamardian, S. [1968], Existence of solutions of certain systems of nonlinear inequalities, Num. Math. 12, 327-334.

Moré, J. [1970], A class of nonlinear functions and the convergence of Gauss-Seidel and Newton-Gauss-Seidel iterations, Ph.D. Dissertation, University of Maryland.

Moré, J., and Rheinboldt, W. [1970], On P- and S-functions and related classes of n-dimensional, nonlinear mappings, Computer Science Center, University of Maryland, Technical Report 70-120.

Ortega, J., and Rheinboldt, W. [1967], Monotone iterations for nonlinear equations with application to Gauss-Seidel methods, SIAM J. Num. Anal. 4, 171-190.

Ortega, J., and Rheinboldt, W. [1970a], Local and global convergence of generalized linear iterations, Studies in Numerical Analysis 2, 122-143, SIAM Publications, Philadelphia, Pa.

Ortega, J., and Rheinboldt, W. [1970b], Iterative Solution of Nonlinear Equations in Several Variables, Academic Press, Inc., New York, N.Y.

Porsching, T. [1969], Jacobi and Gauss-Seidel methods for nonlinear network problems, SIAM J. Num. Anal. 6, 437-449.

Rheinboldt, W. [1969a], Local mapping relations and global implicit function theorems, Trans. Am. Math. Soc. 138, 183-198.

Rheinboldt, W. [1969b], On M-functions and their application to nonlinear Gauss-Seidel iterations and network flows, Ges. f. Math. u. Datenverarb., Birlinghoven/Germany, Tech. Rep. BMwF-GMD-23; J. Math. Anal. Appl., to appear.

Sandberg, I. W., and Willson, Jr., A. N. [1969a], Some theorems on properties of DC equations of nonlinear networks, The Bell System Tech. J. 48, 1-34.

Sandberg, I. W., and Willson, Jr., A. N. [1969b], Some network-theoretic properties of nonlinear DC transistor networks, The Bell System Tech. J. 48, 1293-1311.

Schechter, S. [1962], Iteration methods for nonlinear problems, Trans. Am. Math. Soc. 104, 179-189.

Stiemke, E. [1915], Über positive Lösungen homogener linearer Gleichungen, Math. Annalen 76, 340-342.

Varga, R. [1962], Matrix Iterative Analysis, Prentice Hall, Englewood Cliffs, N.J.

Willson, Jr., A. N. [1968], On the solution of equations for nonlinear resistive networks, The Bell System Tech. J. 48, 1755-1773.